Woody International Publish Limited
An Multidisciplinary Academic Journal Publisher
Journal of Artificial Intelligence and Information, Volume 3, 2026
https://www.woodyinterpub.com/

A Study on the Application of Big Data Technologies in Computer Network Intrusion Detection Systems

Xavier Daimari

Research Scholar, Department of Computer Science and Engineering, Punjabi University, Patiala, India

Abstract: ***With the continuous advancement of technological capabilities, computer and information technologies have undergone sustained development, leading to the emergence of big data technology. Big data technology encompasses a set of techniques for collecting, analyzing, organizing, and storing diverse types of information data. It enables effective management of various information categories and, when applied to computer network intrusion detection, enhances both the accuracy and precision of network security systems while facilitating the automation and intelligent operation of intrusion detection processes. Consequently, this integration holds significant practical value in ensuring the safe and stable operation of computer networks.***

Keywords: Big data technology; Computer network intrusion detection; Application.

1. INTRODUCTION

Network intrusion detection is currently one of the most popular research topics internationally. It is an advanced dynamic network security control technology. However, as an emerging network security vulnerability detection technology, there is a lack of experience and inspiration for reference, and there are still many problems that need to be solved. Apply big data mining techniques to computer network intrusion detection. It can quickly determine the location and time of intrusion, and propose corresponding security defense strategies to ensure the security of computer networks. Therefore, it is particularly important to study the application of big data technology in computer network intrusion detection.

2. OVERVIEW OF BIG DATA TECHNOLOGY

In today's information age, big data has become an inevitable demand and a key technology for promoting China's economic development. Big data is a type of data that can analyze, organize, and store large amounts of information, and its essence is the information technology of information data. At the same time, through information mining of big data, innovation can be made in concepts, models, technologies, and applications, continuously optimizing and innovating big data technology to better meet the needs of the current information age and contribute to the development of today's society. Big data is essentially a database used to store various industries and data information. When relevant parties collect information from databases, reasonable technical means can be used to collect data, thereby improving the timeliness of data collection. Big data is based on the Internet, the Internet of Things, and enterprise data to build data sources, and to collect and store data through extraction, conversion, loading and other methods. On this basis, automated management of various data has been achieved. If users wish to collect relevant data, they can obtain corresponding authorization from the database, making the data visualized. Deng and Yang [1] proposed multi-layer defense strategies and privacy-preserving enhancements to mitigate membership reasoning attacks in federated learning frameworks (Deng & Yang, 2025). In the field of human-computer interaction, Sun [2] explored AI-assisted UI design, demonstrating how generative tools can enhance both efficiency and creativity (Sun, 2026). Liu [3] focused on volatility forecasting and early-warning market stress detection, employing leakage-safe evaluation methods with tree ensembles and transformers (Liu, 2026). Yi [4] addressed fair-exposure ad allocation for small businesses and underserved creators using contextual bandits-with-knapsacks in real-time settings (Yi, 2025). Tang et al. [5] designed and optimized shallow-angle grating couplers for vertical emission from indium phosphide devices, contributing to photonic integration (Tang et al., 2020). Tian, Wang, and Cui [6] improved brain tumor image segmentation by integrating GSConv modules and ECA attention mechanisms into the Unet architecture (Tian et al., 2024). Ximeng and Yiming [7] developed an offline conservative reinforcement learning model for transaction authorization, balancing fraud risk and customer friction (Ximeng & Yiming, 2026). Zhao et al. [8] optimized deep learning models for dynamic market behavior prediction, advancing financial time-series analysis (Zhao et al.,

2025). Yang et al. [9] designed a full-cycle intelligent risk control system for pre-loan, mid-loan, and post-loan lending, enabling AI-driven closed-loop management of online credit security (Yang et al., 2025). Shen et al. [10] applied the whale optimization algorithm to financial payment fraud detection, demonstrating its effectiveness in identifying fraudulent transactions (Shen et al., 2025). Li [11] optimized AI-driven bid pricing models for non-standard automation projects by leveraging historical financial data and machine learning algorithms (Li, 2026). Ren [12] proposed a novel feature fusion-based and complex contextual model for smoking detection, achieving robust performance in real-world scenarios (Ren, 2024). Zhou [13] diagnosed bottlenecks in international automotive sales funnels using gradient boosting trees, providing evidence from cross-regional team efficiency evaluations (Zhou, 2026). Finally, Wensi [14] explored AI-enabled data visualization marketing for automated production lines, focusing on building customer trust and improving lead-to-order conversion rates (Wensi, 2026).

3. OVERVIEW OF NETWORK INTRUSION DETECTION

3.1 Concept of Network Intrusion Detection

To ensure the security of the network, network intrusion detection is an effective means. Network intrusion detection is to check the operation of the network. Based on the computer user's behavior in using the computer, it can determine whether this behavior may pose a risk of intrusion to the network. If the user's behavior may bring a risk of intrusion to the Internet, then network intrusion detection can intercept it and report it to the network user, so it can maximize the security of the network and enable the network to operate smoothly. Network intrusion detection, as an internal system based on the network, is a very meaningful network security management technology. It can collect information from various system sources and analyze the working status of computer networks based on this data, so as to identify whether the computer network is at risk of being invaded or attacked by other hackers. In addition, through network intrusion detection technology, the overall operation of the network can be comprehensively monitored. When the network is working normally, network intrusion detection will be performed until the entire network stops running. This technology automatically collects the network operation status, establishes corresponding logs, and uploads them to the system. While managing the network, relevant personnel can also access relevant network detection records to evaluate the stability of the computer network. The system can also automatically generate corresponding response reports and report to relevant personnel when intrusion behavior is detected. For some threatening network intrusion detection, firewalls can be used to automatically detect areas. When attackers encounter difficult to resist attacks, they need to quickly take corresponding measures upon receiving intrusion information to ensure the safe and stable operation of computer networks.

3.2 Classification of Intrusion Detection Technologies

According to different intrusion detection technologies, their classification can be divided into: anomaly monitoring and misuse behavior recognition; According to the different detection targets, they can be divided into host based, network-based, and host based hybrid types. Intrusion detection mainly consists of the following three parts: first, information collection, collecting device operating status, user behavior and other information, and collecting network protocol and network traffic information. The second is information analysis, which includes statistical analysis of information. When abnormal situations are detected, the system will immediately sound an alarm and send a fault log to the console. Thirdly, the console selects the appropriate processing method based on the type of abnormal information received.

4. CURRENT ISSUES WITH INTRUSION DETECTION SYSTEMS

Intrusion detection is an effective technology to ensure network security, including anomaly detection, feature detection, protocol analysis, state detection, etc. In practical applications, intrusion detection systems usually use several different methods combined together, but there are still many aspects that need further research and improvement.

4.1 High error rate

False and missed reports are the most prominent performance indicators in current intrusion detection systems (IDS). According to statistics, 3000-10000 attackers use vulnerabilities in intrusion detection systems every year, while the current vulnerability detection rate of intrusion detection systems is only 50%. Many attacks are aimed at

invading these weaknesses. Therefore, how to effectively improve the detection rate and reduce the error recognition rate of intrusion detection systems is a key issue that urgently needs to be addressed in current research on intrusion detection systems. Traditional network intrusion detection methods typically involve multiple intrusion detectors simultaneously detecting a target computer, while multiple intrusion detectors can only detect one target computer. The commonly used methods currently involve scanning multiple target hosts, but there are issues such as lengthy scanning times and inability to cover all possible vulnerabilities. At present, the false alarm rate of intrusion detection systems is very high, and the fundamental reason is that the detection accuracy of IDS is not high enough, and existing detection methods have defects. Currently, there is no good solution for large-scale mixed distributed attacks. For example, commonly used statistical methods detect anomalies in networks, but their thresholds are difficult to accurately determine errors. A threshold value that is too small can cause many errors, and a threshold value that is too high can also cause many errors.

4.2 Lack of proactive protection

Intrusion detection technology is a passive and limited technique that cannot effectively perform intrusion detection. In recent years, an increasing number of new technologies such as worms, Trojans, hacker software, etc. have posed a great threat to the security of computer networks. When existing security measures cannot guarantee network security, using IDS to ensure network security is a feasible method. In IDS, existing detection rules can be updated through pre-defined methods or feature descriptions, which are usually lagging. When new vulnerabilities are discovered, there will be methods and means to immediately attack them. However, it will take a long time to discover reasonable detection and defense rules to address this vulnerability. In fact, new hacker technology updates also require a certain amount of time, during which time it is sufficient for hackers to launch attacks.

4.3 Lack of accurate positioning and processing mechanism

IDS can only identify IP addresses and cannot determine the specific data source.
For example, the function of a firewall is to scan network security, identify unsafe factors, filter unsafe traffic, and ensure the normal operation of the network. The main function of an intrusion detection system is to monitor data in the network in real-time, detect abnormal states, and issue timely alarms. Firewalls and IDS are essentially a whole, both aimed at preventing hacker intrusions and protecting network security. There is a significant difference between the two: the role of a firewall is to comprehensively manage the entire network; IDS, on the other hand, monitors the operational status of the network to identify potential security vulnerabilities and provide effective protection against them. Currently, most firewalls and IDS work independently. Both methods have their own limitations that hinder their practical application. When an attack is detected, only a few ports or outputs can be closed, which can have a significant impact on other ordinary users and there is no effective response mechanism.

5. APPLICATION OF DATA MINING TECHNOLOGY IN NETWORK INTRUSION DETECTION

Data mining technology is one of the important components of big data technology. Combining existing research results and work experience, this technology has achieved good results in actual network intrusion. It is as follows:

5.1 System Model and Detection Methods

Applying data mining techniques to network intrusion detection can form a distributed network intrusion detection system based on technologies such as big data and cloud computing. The system uses mobile agents to collect complete detection content and transmit it to the event sequence generator, achieving effective behavior recognition through data mining techniques. This method analyzes the similarity between discovered information or relevant rules, enabling decision-makers to make final decisions and ensuring the security of computer network systems.

Currently, there are two main types of network intrusion detection technologies. One of them is a host based intrusion detection system that targets different operating system characteristics and can quickly detect attacks at the application layer. However, this method requires the host and audit system to work together and does not have real-time capabilities; Another type is network-based intrusion detection, which utilizes data collected from the network to analyze suspicious intrusion detection behaviors. It does not rely on hosts and can operate according to standard network protocols. Through an intelligent network intrusion detection system, the two can be organically

integrated to better adapt to intrusion detection requirements in various environments. When the detection environment changes, only corresponding modifications need to be made to the corresponding data, without the need for complete modifications to hardware, software, and protocols. In the network environment, intelligent intrusion detection systems can quickly identify new technologies and applications, and can expand the rule library through self-learning functions, thereby achieving adaptability to the network.

5.2 System Architecture Design

Currently, two types of network intrusion system architectures are commonly used internationally. One is to build a unified central platform to monitor network intrusions. However, this architecture is only suitable for small-scale network management, and in large-scale network management, there may be issues with inaccurate detected information. The second type of network architecture is a central architecture centered around subnetworks. Each region has intrusion detection professionals, and each system can be seen as an independent system. This system structure can comprehensively and accurately detect each subnet, which is more in line with the requirements of network intrusion detection system structure. When designing a network IDS architecture, it is necessary to collect preprocessed data sources and import them into a data warehouse. Then, using data mining techniques, a data mining engine is constructed, and the data information is transmitted to both the detection module and the rule base. At the same time, the analyzed and processed data from the rule base is transmitted to the detection module. Then the detection module checks the data to determine if there is a network intrusion. On this basis, when network intrusion behavior is detected, relevant information must be transmitted to the corresponding intrusion module to achieve the purpose of detection. As shown in Figure 1, data mining techniques.

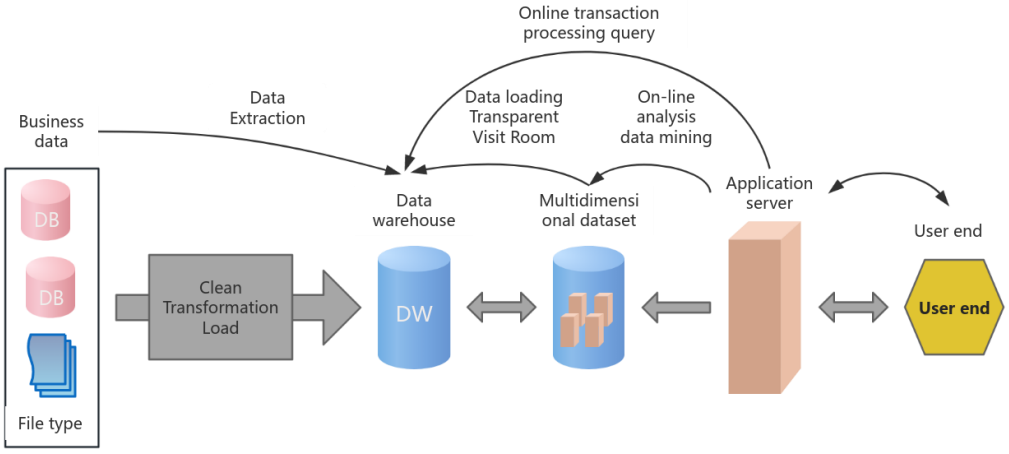

Figure 1: Data Mining Techniques

5.3 Application of Data Mining Algorithms in Network Intrusion Detection

Association rules refer to the display of relationships and rules between a series of objects in a set of data. When data mining algorithms discover network vulnerabilities, there is a correlation between program execution and user behavior, which is often reflected in the associated dataset. By using association analysis methods, it is possible to quickly obtain the correlation between multiple data elements, and then determine the connections between each element. At the same time, sequence analysis techniques can be used to describe the association between multiple transactions, extract sequence patterns from transactions, and meet users' minimum needs for frequent sequences.

In the audit process, it is usually not possible to determine in advance the data required by users, but the following two methods can be used to analyze the rules that users are concerned about. One is to first obtain suitable rules, and then sort each rule according to the user's level of interest, removing unnecessary rules; The second is to utilize existing knowledge, use the rules to be processed as conditional constraints, and then conduct mining. The basic mechanism of using association rules for network intrusion detection is to use relevant rule algorithms to quickly identify various unknown intrusion patterns, in order to achieve monitoring and prevention of network algorithms and improve network security and stability. Compared with other network intrusion detection methods, the

algorithm using association rules can quickly and efficiently mine some unknown network intrusion patterns, display various information and behavior patterns of the current computer user, and verify the current behavior pattern against historical data and behavior. If the difference between the two is too large, it can be judged as an intrusion.

6. CONCLUSION

In summary, using big data mining technology for intrusion detection in computer networks can effectively enhance their intelligence level and improve their work efficiency and quality. In addition, with the development of China's economy and society, the demand for network security is also increasing. Therefore, in order to effectively improve the efficiency and quality of computer network intrusion system operation, relevant departments need to increase research on big data technology and continuously improve and optimize it. In addition, when implementing computer network intrusion detection, relevant personnel also need to have a full understanding and analysis of the application of big data technology to ensure its successful implementation.

REFERENCES

[1] Deng, X., & Yang, J. (2025, August). Multi-Layer Defense Strategies and Privacy Preserving Enhancements for Membership Reasoning Attacks in a Federated Learning Framework. In 2025 5th International Conference on Computer Science and Blockchain (CCSB) (pp. 278-282). IEEE.
[2] Sun, Lingxin. "AI-Assisted UI Design: Enhancing Efficiency and Creativity through Generative Tools." Journal of Computer Technology and Applied Mathematics 3.1 (2026): 19-27.
[3] Liu, Ting. "Volatility Forecasting and Early-Warning Market Stress Detection: A Leakage-Safe Evaluation with Tree Ensembles and Transformers." (2026).
[4] Yi, X. (2025, October). Real-Time Fair-Exposure Ad Allocation for SMBs and Underserved Creators via Contextual Bandits-with-Knapsacks. In Proceedings of the 2025 2nd International Conference on Digital Economy and Computer Science (pp. 1602-1607).
[5] Tang, Y., Kojima, K., Gotoda, M., Nishikawa, S., Hayashi, S., Koike-Akino, T., ... & Klamkin, J. (2020). Design and Optimization of Shallow-Angle Grating Coupler for Vertical Emission from Indium Phosphide Devices.
[6] Tian, Q., Wang, Z., & Cui, X. (2024). Improved Unet brain tumor image segmentation based on GSConv module and ECA attention mechanism. arXiv preprint arXiv:2409.13626.
[7] Ximeng, Y., & Yiming, Z. (2026). Offline Conservative RL for Transaction Authorization: Smartly Balancing Fraud Risk and Customer Friction. Journal of Economic Theory and Business Management, 3(1), 1-9.
[8] Zhao, S., Lin, Y., Yang, X., Lu, Q., Xue, H., & Jiang, G. (2025). Optimization of Deep Learning Models for Dynamic Market Behavior Prediction. arXiv preprint arXiv:2511.19090.
[9] Yang, X., Xue, H., Hu, Q., & Zhang, Y. (2025, October). Design of a full-cycle intelligent risk control system for pre-loan, mid-loan, and post-loan lending: AI-driven closed-loop management of online credit security. In Proceedings of the 2025 2nd International Conference on Digital Economy and Computer Science (pp. 1022-1027).
[10] Shen, Zepeng, et al. "Research on Application of Whale Optimization Algorithm in Financial Payment Fraud Detection." 2025 4th International Conference on Artificial Intelligence, Internet and Digital Economy (ICAID). IEEE, 2025.
[11] Li, W. (2026). Optimizing AI-Driven Bid Pricing Models for Non-Standard Automation Projects: Leveraging Historical Financial Data and Machine Learning Algorithms.
[12] Z. Ren, "A Novel Feature Fusion-Based and Complex Contextual Model for Smoking Detection," 2024 6th International Conference on Communications, Information System and Computer Engineering (CISCE), Guangzhou, China, 2024, pp. 1181-1185, doi: 10.1109/CISCE62493.2024.10653351.
[13] Zhou, Z. (2026). Bottleneck Diagnosis in International Automotive Sales Funnels Using Gradient Boosting Trees: Evidence from Cross-Regional Team Efficiency Evaluation. Journal of Computer Technology and Applied Mathematics, 3(1), 11-18.
[14] Wensi, L. (2026). AI-Enabled Data Visualization Marketing for Automated Production Lines: Building Customer Trust and Improving Lead-to-Order Conversion. Academic Journal of Natural Science, 3(1), 8-13.